United States Patent and Trademark Office

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/392,170	09/08/1999	MONIKA R. HENZINGER	3792(PD-595)	9654

758   7590   06/17/2003

FENWICK & WEST LLP
SILICON VALLEY CENTER
801 CALIFORNIA STREET
MOUNTAIN VIEW, CA   94041

EXAMINER
ROMERO, ALMARI DEL CARMEN

ART UNIT	PAPER NUMBER
2176	

DATE MAILED: 06/17/2003

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No. 09/392,170	Applicant(s) HENZINGER ET AL.
	Examiner Almari Romero	Art Unit 2176

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE _3_ MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) ☒ Responsive to communication(s) filed on _17 March 2003_ .

2a) ☐ This action is **FINAL**.   2b) ☒ This action is non-final.

3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) ☒ Claim(s) _2-28 and 30-59_ is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) ☐ Claim(s) _____ is/are allowed.

6) ☒ Claim(s) _2-28 and 30-59_ is/are rejected.

7) ☐ Claim(s) _____ is/are objected to.

8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) ☐ The specification is objected to by the Examiner.

10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

11) ☐ The proposed drawing correction filed on _____ is: a) ☐ approved b) ☐ disapproved by the Examiner.

    If approved, corrected drawings are required in reply to this Office action.

12) ☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

13) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a) ☐ All b) ☐ Some * c) ☐ None of:

      1. ☐ Certified copies of the priority documents have been received.

      2. ☐ Certified copies of the priority documents have been received in Application No. _____ .

      3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

14) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).

    a) ☐ The translation of the foreign language provisional application has been received.

15) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☒ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _6_ .

4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .

5) ☐ Notice of Informal Patent Application (PTO-152)

6) ☐ Other: .

DETAILED ACTION

1.   This action is responsive to communications: Amendment filed on 3/17/03 and the IDS

filed on 03/17/03.

2.   The objection to the disclosure containing embedded hyperlinks has been withdrawn as

necessitated by amendment.

3.   The rejection of claims 2-3, 15, 24-25, 28, 30-31, 43, 52-53, and 56 under 35 U.S.C. 112,

second paragraph, has being indefinite as been withdrawn as necessitated by amendment.

4.   The rejection of claims 7,25,28, and 35 under 35 U.S.C. 112, second paragraph, as being

indefinite has been withdrawn as necessitated by amendment.

5.   The rejection of claims 23 and 51 under 35 U.S.C. 112, second paragraph, as being

indefinite has been withdrawn as necessitated by amendment.

6.   The rejection of claims 1, 4-5, 29, 32-33, and 57 under 35 U.S.C. 102(e) as being

anticipated by Page has been withdrawn as necessitated by amendment.

7.   The rejection of claims 13, 18-20, 41, 46-48, and 58-59 under 35 U.S.C. 102(e) as being

anticipated by Bharat et al. as been withdrawn as necessitated by amendment.

8.   The rejection of claims 2-3, 27, 30-31, and 55 under 35 U.S.C. 103(a) as being

unpatentable over Page and Singhal has been withdrawn as necessitated by amendment.

9.   The rejection of claims 6 and 34 under 35 U.S.C. 103(a) as being unpatentable over Page

and Bharat has been withdrawn as necessitated by amendment.

10.   The rejection of claims 7-12 and 35-40 under 35 U.S.C. 103(a) as being unpatentable

over Singhal and Bharat has been withdrawn as necessitated by amendment.

11.   The rejection of claims 14, 21-23, 26, 42, 49-51, and 54 under 35 U.S.C. 103(a) as being

unpatentable over Bharat and Page has been withdrawn as necessitated by amendment.

12.   The rejection of claims 15-17, 24-25, 28, 43-45, 52-53, and 56 under 35 U.S.C. 103(a) as

being unpatentable over Bharat, Page, and Singhal has been withdrawn as necessitated by

amendment.

13.   Claims 1 and 29 have been cancelled and claims 2-28 and 30-59 are pending in the case.

Claims 2, 7, 9, 13, 15, 20, 21, 24, 25, 27, 28, 30, 35, 37, 41, 43, 48, 49, 52, 53, 55, 56, 57, 58,

and 59 are independent claims.

Information Disclosure Statement

14.   The reference Bray, T. "Measuring the Web" in the information disclosure statement

(IDS) re-submitted on 3/17/03 has been considered by the Examiner.

Claim Rejections - 35 USC § 112

15.   The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the
> subject matter which the applicant regards as his invention.

16.   Claims 16, 44, and 56 are rejected under 35 U.S.C. 112, second paragraph, as being

indefinite for failing to particularly point out and distinctly claim the subject matter which

applicant regards as the invention. The limitations "adding the selected host to the host set " and

"adding the selected document to the document set of the selected host", it is unclear of why a

host or document is being added if it already exist in the host set or document set for selection.

Claim Rejections - 35 USC § 103

17.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in
> section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are
> such that the subject matter as a whole would have been obvious at the time the invention was made to a person
> having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the
> manner in which the invention was made.

18.   **Claims 2, 4-9, 11-12, 15-19, 24-25, 27-28, 30, 32-37, 43-47, and 52-57 are rejected**

under 35 U.S.C. 103(a) as being unpatentable over Pitkow et al. (USPN 6,457,028 B1- filed

on 09/1999) in view of Singhal (USPN 6,370,527 B1 – filed on 12/1998).

Regarding independent claims 2, 27, 30, 55, and 57, Pitkow discloses:

A computer-implemented method and computer program product for randomly walking through

a hypertext-linked document set comprising a plurality of documents, wherein at least a subset of

the documents contain a plurality of links to other documents, each document being associated

with a host (Pitkow on col. 7, lines 49-62: teaches web walker for linked documents), the method

comprising:

a) selecting a host; b) selecting at random a document associated with the host; c)

retrieving the selected document d) responsive to occurrence of a random event: d.1) selecting at

random a link in the retrieved document; d.2) selecting at random a document associated with the

host; and d.3) retrieving the selected document (Pitkow on col. 7, lines 49-62: teaches a web

walker that automatically follows links on a document and collects the linked documents; the

document collection could be a randomly selected collection of documents and on col. 10, lines

15-31: teaches the collection may be gathered through a "web crawl" where linked pages are

obtained, or through a query to one or more search engines; the document collection is identified based on the URL address (host));

e) responsive to non-occurrence of the random event: e.1) selecting at random a link in the retrieved document; and e.2) retrieving a document referenced by the selected link; and f) and f) repeating d) and e) until a predetermined condition is met (Pitkow on col. 7, lines 49-62 and col. 10, lines 15-31: teaches a collection of linked documents can be randomly selected and can be gathered through a "web crawl" where linked pages are obtained or through a query to one or more search engines; until a list of web sites along with indicators of corresponding web pages are obtained).

However, Pitkow does not explicitly disclose "selecting at random a host from among the previously selected hosts".

Singhal on col. 1, lines 30-62 and col. 7, lines 21-30: teaches submitting a query among a plurality of search engine devices and selecting search engine device for retrieval of sources.

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Singhal into Pitkow to provide selecting search engine device from a plurality of search engine devices for the retrieval of sources, as taught by Singhal, incorporated into the "web crawl" or one or more search engines, as taught by Pitkow, in order to allow a user to search all of the available portions of a distributed network without having to repeatly reenter their search query.

Regarding dependent claims 4 and 32, Pitkow discloses:

wherein the document set is the World Wide Web, and wherein each document is a web

page (Pitkow on col. 7, lines 15-31: teaches document collection can be a particular web site, a

set of web sites, or event the entire Web itself).

Regarding dependent claims 5 and 33, Pitkow discloses:

wherein each host corresponds to a domain (Pitkow on col. 10, lines 15-31: teaches

website found in URL address "www.companyabc.com" (domain)).

Regarding dependent claims 6 and 34, Pitkow discloses:

performing a second two-level random walk through the hypertext-linked document set

(Pitkow on col. 7, lines 15-31 and on col. 10, lines 15-60: teaches a "web walker" and randomly

selecting and collecting linked documents; the collection may be gathered based on sampling of

web pages, through a "web crawl" where linked pages are obtained, or through a query to one or

more search engines; a list of web sites is obtained and further processing of ranking in-links of

the list of web pages or web sites is performed).

Pitkow does disclose "a second two-level random walk" on col. 7, lines 49-62: teaches

web walker is used for following links and collecting linked documents which these documents

can be randomly selected; another process is performed by determining the frequency of linkage

for each document within the document collection and on col. 10, lines 15-60: teaches a query to

one or more search engines, in other words, each search engine such as a "web crawl" can

obtained linked pages and can request other search engine to obtain more linked pages (as a

second random walk).

It would have been obvious to a person of ordinary skill in the art at the time the

invention was made to have modified Pitkow to provide a way for randomly selecting linked

documents to create a document collection incorporated into a search engine to contain the

document collection for other search engines to query to obtain more linked pages which will aid

the user in finding the desired information.

Regarding independent claims 7 and 35, Pitkow discloses:

A computer-implemented method for randomly walking through a hypertext-linked document set

comprising a plurality of documents, wherein at least a subset of the documents contain a

plurality of links to other documents, each document being associated with a host (Pitkow on col.

7, lines 49-62: teaches web walker for linked documents), the method comprising:

b) initializing a document set for each host in the host set; d) selecting at random a

document from the document set of the selected host; e) responsive to the selected document

containing at least one link: e.1) selecting at random a link from the selected document; e.2)

selecting a document corresponding to the selected link e.3) selecting a host corresponding to the

selected document (Pitkow on col. 7, lines 49-62: teaches a web walker that automatically

follows links on a document (for the selection at random a link) and collects the linked

documents; the document collection (document set) could be a randomly selected collection of

documents (selected documents) and on col. 10, lines 15-31: teaches the collection may be

gathered through a "web crawl" where linked pages are obtained, or through a query to one or

more search engines; each search engine can have a document collection or a list of web pages or

web sites for other search engines to perform queries).

However, Pitkow does not explicitly disclose "a) initializing a host set; c) selecting at

random a host from the host set; e.4) adding the selected host to the host set; e.5) adding the

selected document to the document set of the selected host".

Singhal discloses "a) initializing a host set" on col. 1, lines 30-62: teaches plurality of search engine devices; "c) selecting at random a host from the host set" on col. 1, lines 30-32 and col. 7, lines 21-30: teaches submitting a query among a plurality of search engine devices (host set) and selecting search engine device for retrieval of sources; "e.4) adding the selected host to the host set" on col. 2, lines 18-27: teaches a number of search engines that found the same source can be added based on availability; "e.5) adding the selected document to the document set of the selected host" on col. 4, lines 1-17: teaches source such as web page or document can be found by the search engine device and added in search result).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Singhal into Pitkow to provide selecting search engine device from a plurality of search engine devices (host set) for the retrieval of sources, as taught by Singhal, incorporated into the "web crawl" or one or more search engines, as taught by Pitkow, in order to allow a user to search all of the available portions of a distributed network without having to repeatly reenter their search query.

Regarding dependent claims 8 and 36, Singhal discloses:

is performed responsive to the selected host not being in the host set; and is performed responsive to the selected document not being in the document set of the selected host (Singhal col. 2, lines 18-27: teaches number of search engines can be selected or not selected based on availability and if being able to fine the requested source (document)).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Singhal into Pitkow to provide selecting search engine device from a plurality of search engine devices for the retrieval of sources, as taught by Singhal,

incorporated into the "web crawl" or one or more search engines, as taught by Pitkow, in order to allow a user to search all of the available portions of a distributed network without having to repeatly reenter their search query.

Regarding independent claims 9 and 37, Singhal discloses:

A computer-implemented method for randomly walking through a hypertext-linked document set comprising a plurality of documents, wherein at least a subset of the documents contain a plurality of links to other documents, each document being associated with a host (Pitkow on col. 7, lines 49-62: teaches web walker for linked documents), the method comprising:

b) initializing a document set for each host in the host set; d) selecting at random a document from the document set of the selected host; e) responsive to non-occurrence of the random event, and further responsive to the selected document containing at least one link: e.1) selecting at random a link from the selected document; e.2) selecting a document corresponding to the selected link e.3) selecting a host corresponding to the selected document (Pitkow on col. 7, lines 49-62 and col. 10, lines 15-31: teaches a collection of linked documents (document set) can be randomly selected (selecting at random a document) and can be gathered through a "web crawl" where linked pages are obtained or through a query to one or more search engines; until a list of web sites along with indicators of corresponding web pages are obtained; in other words, the document can be selected among one or more search engines containing a gathered document collection or a list of web pages or web sites).

Singhal discloses "a) initializing a host set" on col. 1, lines 30-62: teaches plurality of search engine devices; "c) selecting at random a host from the host set" on col. 1, lines 30-32 and col. 7, lines 21-30: teaches submitting a query among a plurality of search engine devices (host

set) and selecting search engine device for retrieval of sources; "e.4) adding the selected host to

the host set" on col. 2, lines 18-27: teaches a number of search engines that found the same

source can be added based on availability; "e.5) adding the selected document to the document

set of the selected host" on col. 4, lines 1-17: teaches source such as web page or document can

be found by the search engine device and added in search result).

It would have been obvious to a person of ordinary skill in the art at the time the

invention was made to have modified Singhal into Pitkow to provide selecting search engine

device from a plurality of search engine devices (host set) for the retrieval of sources, as taught

by Singhal, incorporated into the "web crawl" or one or more search engines, as taught by

Pitkow, in order to allow a user to search all of the available portions of a distributed network

without having to repeatly reenter their search query.

Regarding dependent claims 11 and 39, Pitkow discloses:

wherein the document set is the World Wide Web, and wherein each document is a web

page (Pitkow on col. 7, lines 15-31: teaches document collection can be a particular web site, a

set of web sites, or event the entire Web itself).

Regarding dependent claims 12 and 40, Pitkow discloses:

wherein each host corresponds to a domain (Pitkow on col. 10, lines 15-31: teaches

website found in URL address "www.companyabc.com" (domain)).

Regarding independent claims 15 and 43, Pitkow discloses:

A computer-implemented method for measuring relative quality of a search engine index

(Pitkow on col. 10, lines 54-60: teaches analyzing citation index for each web page within a web

site), comprising:

a) performing a two-level random walk among documents within a document set; a.1)

selecting a host ; a.2) selecting at random a document associated with the host; a.3) retrieving the

selected document; a.3.1) responsive to occurrence of a random event; a.3.1.2) retrieving the

selected document (Pitkow on col. 7, lines 15-31 and on col. 10, lines 15-60: teaches a "web

walker" and randomly selecting and collecting linked documents (selecting at random a

document); the collection may be gathered based on sampling of web pages, through a "web

crawl" where linked pages are obtained, or through a query to one or more search engines

(selecting a host); a list of web sites is obtained and further processing of ranking in-links of the

list of web pages or web sites is performed (two-level random walk));

a.3.2) responsive to non-occurrence of the random event:a.4) selecting at random a link in

the retrieved document; and a.5) retrieving a document referenced by the selected link (Pitkow

on col. 7, lines 49-62 and col. 10, lines 15-31: teaches a collection of linked documents can be

randomly selected and can be gathered through a "web crawl" where linked pages are obtained

or through a query to one or more search engines; until a list of web sites along with indicators of

corresponding web pages are obtained);

b) for each document encountered in the random walk, determining whether the

document is indexed by the search engine index (Pitkow on col. 10, lines 54-60: teaches

constructing and analyzing citation index for each web page; a citation index is a listing of all the

links contained in the page; wherein for each pair of web sites would reveal the number of times

each of the sites are both cited by the same page and the same site); and c) aggregating the

results of b) (Pitkow on col. 11, lines 10-15: teaches aggregate all the links to a "destination"

web site).

Pitkow does disclose "a second two-level random walk" on col. 7, lines 49-62: teaches web walker is used for following links and collecting linked documents which these documents can be randomly selected; another process is performed by determining the frequency of linkage for each document within the document collection and on col. 10, lines 15-60: teaches a query to one or more search engines, in other words, each search engine such as a "web crawl" can obtained linked pages and can request other search engine to obtain more linked pages (as a second random walk).

However, Pitkow does not explicitly disclose "selecting at random a host from among the previously selected hosts".

Singhal on col. 1, lines 30-62 and col. 7, lines 21-30: teaches submitting a query among a plurality of search engine devices and selecting search engine device for retrieval of sources.

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Singhal into Pitkow to provide selecting search engine device from a plurality of search engine devices for the retrieval of sources, as taught by Singhal, incorporated into the "web crawl" or one or more search engines, as taught by Pitkow, in order to allow a user to search all of the available portions of a distributed network without having to repeatly reenter their search query.

Regarding dependent claims 16 and 44, the limitations of claims 16 and 44 are similar to those in rejecting claim 7 and are rejected under the same rationale.

Regarding dependent claims 17 and 45, the limitations of claims 17 and 45 are similar to those in rejecting claim 8 and are rejected under the same rationale.

Regarding dependent claims 18 and 46, Singhal discloses:

wherein each document contains a plurality of words, and wherein b) comprises, for each document encountered in the random walk: b.1) selecting at least one word from the document; b.2) performing a query on the search engine index based on the selected at least one word, to obtain search results; and b.3) determining whether the document is included in the obtained search results (Singhal on col. 2, lines 42-48: teaches search terms (words) found in resources (document) for a query to performed by the search engine; weighting the results obtained from the search engines).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Singhal into Pitkow to provide a plurality of search terms (words) for performing a query by the search engine devices, as taught by Singhal, incorporated into the search engine "web crawl" or one or more search engines, as taught by Pitkow, in order to allow a user to search all of the available portions of a distributed network without having to repeatly reenter their search query.

Regarding dependent claims 19 and 47, Singhal discloses:

wherein b.1) comprises selecting at least one word based on rarity (Singhal on col. 2, lines 42-48: teaches search terms (words) found in resources (document) and ranking the occurrence of search terms (determine if term occurs more than once)).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Singhal into Pitkow to provide a plurality of search terms (words) for performing a query by the search engine devices, as taught by Singhal, incorporated into the search engine "web crawl" or one or more search engines, as taught by Pitkow, in order

to allow a user to search all of the available portions of a distributed network without having to

repeatly reenter their search query.

Regarding independent claims 24 and 52, Pitkow discloses:

A computer-implemented method for measuring relative quality of a target document in a

document set comprising a plurality of documents, wherein at least a subset of the documents

contain a plurality of links to other documents, wherein each document is associated with a host

(Pitkow on col. 10, lines 54-60: teaches analyzing citation index for each web page within a web

site (document collection)), the method comprising:

a) performing a two-level random walk among documents within a document set, by: a.1)

selecting a host; a.2) selecting at random a document associated with the host ; a.3) retrieving the

selected document; a.4) responsive to occurrence of a random event: a.4.2) selecting at random a

document associated with the host; and a.4.3) retrieving the selected document (Pitkow on col.

7, lines 15-31 and on col. 10, lines 15-60: teaches a "web walker" and randomly selecting and

collecting linked documents (selecting at random a document); the collection may be gathered

based on sampling of web pages, through a "web crawl" where linked pages are obtained, or

through a query to one or more search engines (selecting a host); a list of web sites is obtained

and further processing of ranking in-links of the list of web pages or web sites is performed (two-

level random walk));

a.5) responsive to non-occurrence of the random event: a.5.1) selecting at random

a link in the retrieved document; and a.5.2) retrieving a document referenced by the selected link

(Pitkow on col. 7, lines 49-62 and col. 10, lines 15-31: teaches a collection of linked documents

can be randomly selected and can be gathered through a "web crawl" where linked pages are

obtained or through a query to one or more search engines; until a list of web sites along with

indicators of corresponding web pages are obtained);

b) determining a quality metric responsive to the number of documents encountered

during the two-level random walk that link to the target document (Pitkow on col. 10, lines 54-

60: teaches constructing and analyzing citation index for each web page; a citation index is a

listing of all the links contained in the page; wherein for each pair of web sites would reveal the

number of times each of the sites are both cited by the same page and the same site (determining

quality of the site with collection of documents)).

Pitkow does disclose "a second two-level random walk" on col. 7, lines 49-62: teaches

web walker is used for following links and collecting linked documents which these documents

can be randomly selected; another process is performed by determining the frequency of linkage

for each document within the document collection and on col. 10, lines 15-60: teaches a query to

one or more search engines, in other words, each search engine such as a "web crawl" can

obtained linked pages and can request other search engine to obtain more linked pages (as a

second random walk).

However, Pitkow does not explicitly disclose "selecting at random a host from among the

previously selected hosts".

Singhal on col. 1, lines 30-62 and col. 7, lines 21-30: teaches submitting a query among a

plurality of search engine devices and selecting search engine device for retrieval of sources.

It would have been obvious to a person of ordinary skill in the art at the time the

invention was made to have modified Singhal into Pitkow to provide selecting search engine

device from a plurality of search engine devices for the retrieval of sources, as taught by Singhal,

incorporated into the "web crawl" or one or more search engines, as taught by Pitkow, in order to

allow a user to search all of the available portions of a distributed network without having to

repeatly reenter their search query.

Regarding independent claims 25 and 53, Pitkow discloses:

A computer-implemented method for measuring relative quality of a target document in a

document set comprising a plurality of documents, wherein at least a subset of the documents

contain a plurality of links to other documents, wherein each document is associated with a host,

(Pitkow on col. 10, lines 54-60: teaches analyzing citation index for each web page within a web

site (document collection)), the method comprising:

a) performing a two-level random walk among documents within a document set, by:

initializing a document set for each host in the host set; a.4) responsive to occurrence of a

random event ((Pitkow on col. 7, lines 15-31 and on col. 10, lines 15-60: teaches a "web walker"

and randomly selecting and collecting linked documents; the collection (document set) may be

gathered based on sampling of web pages, through a "web crawl" where linked pages are

obtained, or through a query to one or more search engines (selecting a host); a list of web sites

is obtained and further processing of ranking in-links of the list of web pages or web sites is

performed (two-level random walk))

a.5) responsive to non-occurrence of the random event: a.5.1) selecting at random

a document from the document set of the selected host; a.5.2) responsive to the selected

document containing at least one link, a.5.2.1) selecting at random a link from the selected

document; a.5.2.2) selecting a document corresponding to the selected link; and a.5.2.3) selecting

a host corresponding to the selected document (Pitkow on col. 7, lines 49-62 and col. 10, lines

15-31: teaches a collection of linked documents can be randomly selected (selecting at random a document) and can be gathered through a "web crawl" where linked pages are obtained or through a query to one or more search engines; until a list of web sites along with indicators of corresponding web pages are obtained; in other words, the document can be selected from a plurality of search engines also containing document collections or list of web pages or web sites);

        b) determining a quality metric responsive to the number of documents encountered during the two-level random walk that link to the target document (Pitkow on col. 10, lines 54-60: teaches constructing and analyzing citation index for each web page; a citation index is a listing of all the links contained in the page; wherein for each pair of web sites would reveal the number of times each of the sites are both cited by the same page and the same site (determining quality of the site with collection of documents)).

Pitkow does disclose "a second two-level random walk" on col. 7, lines 49-62: teaches web walker is used for following links and collecting linked documents which these documents can be randomly selected; another process is performed by determining the frequency of linkage for each document within the document collection and on col. 10, lines 15-60: teaches a query to one or more search engines, in other words, each search engine such as a "web crawl" can obtained linked pages and can request other search engine to obtain more linked pages (as a second random walk).

However, Pitkow does not explicitly disclose "selecting at random a host from among the previously selected hosts"; "a) initializing a host set; c) selecting at random a host from the host

set; e.4) adding the selected host to the host set; e.5) adding the selected document to the document set of the selected host".

Singhal does disclose "selecting at random a host from among the previously selected hosts " on col. 1, lines 30-62 and col. 7, lines 21-30: teaches submitting a query among a plurality of search engine devices and selecting search engine device for retrieval of sources; "a) initializing a host set" on col. 1, lines 30-62: teaches plurality of search engine devices; "c) selecting at random a host from the host set" on col. 1, lines 30-32 and col. 7, lines 21-30: teaches submitting a query among a plurality of search engine devices (host set) and selecting search engine device for retrieval of sources; "e.4) adding the selected host to the host set" on col. 2, lines 18-27: teaches a number of search engines that found the same source can be added based on availability; "e.5) adding the selected document to the document set of the selected host" on col. 4, lines 1-17: teaches source such as web page or document can be found by the search engine device and added in search result).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Singhal into Pitkow to provide selecting search engine device from a plurality of search engine devices (host set) for the retrieval of sources, as taught by Singhal, incorporated into the "web crawl" or one or more search engines, as taught by Pitkow, in order to allow a user to search all of the available portions of a distributed network without having to repeatly reenter their search query.

Regarding claims 28 and 56, the limitations of claims 28 and 56 are similar to those in rejecting claim 25 and are rejected under the same rationale.

19.   **Claims 3, 10, 31, and 38 are rejected under 35 U.S.C. 103(a) as being unpatentable over Pitkow and Singhal, as applied to claims 2, 4-9, 11-12, 15-19, 24-25, 27-28, 30, 32-37, 43-47, and 52-57 above, and in further view of Page (USPN 6,285,999 B1 – filed on 1/1998). Regarding dependent claims 3, 10, 31, and 38,** Pitkow and Singhal disclose the invention substantially as claimed as described *supra.* However, Pitkow and Singhal do not explicitly disclose "a generated random number falling within the predetermined range".

Page does disclose "generated random number falling within the predetermined range" on col. 5, lines 21-59 and col. 6, lines 12-60: teaches pages with no links cause some complication by adding huge amounts to the "random jump factor" during the ranking process, in other words, the "random jump factor" is a factor of the number of jumps from one link of a document to another link.

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Page into Pitkow and Singhal to provide a way to determine the random jump factor during ranking process, as taught by Page, incorporated into the ranking process of Pitkow and Singhal in order for the ranking method to provide superior results.

20.   **Claims 13-14, 20-22, 26, 41-42, 48-50, 54, and 58-59 are rejected under 35 U.S.C. 103(a) as being unpatentable over Pitkow et al. (USPN 6,457,028 B1- filed on 09/1999). Regarding independent claims 13, 41, and 58,** Pitkow discloses:

A computer-implemented method for measuring relative quality of a search engine index (Pitkow on col. 10, lines 54-60: teaches analyzing citation index for each web page within a web site (document collection)), comprising:

a) performing a two-level random walk among documents within a document set (Pitkow on col. 7, lines 15-31 and on col. 10, lines 15-60: teaches a "web walker" and randomly selecting and collecting linked documents; the collection may be gathered based on sampling of web pages, through a "web crawl" where linked pages are obtained, or through a query to one or more search engines; a list of web sites is obtained and further processing of ranking in-links of the list of web pages or web sites is performed);

b) for each document encountered in the random walk, determining whether the document is indexed by the search engine index (Pitkow on col. 10, lines 54-60: teaches constructing and analyzing citation index for each web page; a citation index is a listing of all the links contained in the page; wherein for each pair of web sites would reveal the number of times each of the sites are both cited by the same page and the same site); and c) aggregating the results of b) (Pitkow on col. 11, lines 10-15: teaches aggregate all the links to a "destination" web site).

Pitkow does disclose "a second two-level random walk" on col. 7, lines 49-62: teaches web walker is used for following links and collecting linked documents which these documents can be randomly selected; another process is performed by determining the frequency of linkage for each document within the document collection and on col. 10, lines 15-60: teaches a query to one or more search engines, in other words, each search engine such as a "web crawl" can

obtained linked pages and can request other search engine to obtain more linked pages (as a second random walk).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Pitkow to provide a way for randomly selecting linked documents to create a document collection incorporated into a search engine to contain the document collection for other search engines to query to obtain more linked pages which will aid the user in finding the desired information.

Regarding dependent claims 14 and 42, Pitkow discloses:

a.1) selecting a host; a.2) selecting at random a document associated with the host; a.3) retrieving the selected document; a.4) selecting at random a link in the retrieved document (Pitkow on col. 7, lines 49-62: teaches a web walker that automatically follows links on a document and collects the linked documents; the document collection could be a randomly selected collection of documents and on col. 10, lines 15-31: teaches the collection may be gathered through a "web crawl" where linked pages are obtained, or through a query to one or more search engines; the document collection is identified based on the URL address (host));

a.5) retrieving a document referenced by the selected link (Pitkow on col. 7, lines 49-62 and col. 10, lines 15-31: teaches a collection of linked documents can be randomly selected and can be gathered (retrieving a document) through a "web crawl" where linked pages are obtained (retrieving a document) or through a query to one or more search engines; until a list of web sites along with indicators of corresponding web pages are obtained).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Pitkow to provide a way for randomly selecting linked

documents to create a document collection incorporated into a search engine to contain the document collection for other search engines to query to obtain more linked pages which will aid the user in finding the desired information.

Regarding independent claims 20, 48, and 59, Pitkow discloses:

A computer-implemented method for measuring relative quality of a target document in a document set (Pitkow on col. 10, lines 54-60: teaches analyzing citation index for each web page within a web site (document collection)), comprising:

a) performing a two-level random walk among documents within a document set (Pitkow on col. 7, lines 15-31 and on col. 10, lines 15-60: teaches a "web walker" and randomly selecting and collecting linked documents; the collection may be gathered based on sampling of web pages, through a "web crawl" where linked pages are obtained, or through a query to one or more search engines; a list of web sites is obtained and further processing of ranking in-links of the list of web pages or web sites is performed); and

b) determining a quality metric responsive to the number of times the target document is encountered in the random walk (Pitkow on col. 10, lines 54-60: teaches constructing and analyzing citation index for each web page; a citation index is a listing of all the links contained in the page; wherein for each pair of web sites would reveal the number of times each of the sites are both cited by the same page and the same site (determining quality of the site with collection of documents)).

Pitkow does disclose "a second two-level random walk" on col. 7, lines 49-62: teaches web walker is used for following links and collecting linked documents which these documents can be randomly selected; another process is performed by determining the frequency of linkage

for each document within the document collection and on col. 10, lines 15-60: teaches a query to one or more search engines, in other words, each search engine such as a "web crawl" can obtained linked pages and can request other search engine to obtain more linked pages (as a second random walk).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Pitkow to provide a way for randomly selecting linked documents to create a document collection incorporated into a search engine to contain the document collection for other search engines to query to obtain more linked pages which will aid the user in finding the desired information.

Regarding independent claims 21 and 49, Pitkow discloses:

A computer-implemented method for measuring relative quality of a target document in a document set comprising a plurality of documents, wherein at least a subset of the documents contain a plurality of links to other documents (Pitkow on col. 10, lines 54-60: teaches analyzing citation index for each web page within a web site (document collection)), the method comprising:

a) performing a two-level random walk among documents within a document set (Pitkow on col. 7, lines 15-31 and on col. 10, lines 15-60: teaches a "web walker" and randomly selecting and collecting linked documents; the collection may be gathered based on sampling of web pages, through a "web crawl" where linked pages are obtained, or through a query to one or more search engines; a list of web sites is obtained and further processing of ranking in-links of the list of web pages or web sites is performed); and

b) determining a quality metric responsive to the number of documents encountered during the two-level random walk that link to the target document (Pitkow on col. 10, lines 54-60: teaches constructing and analyzing citation index for each web page; a citation index is a listing of all the links contained in the page; wherein for each pair of web sites would reveal the number of times each of the sites are both cited by the same page and the same site (determining quality of the site with collection of documents)).

Pitkow does disclose "a second two-level random walk" on col. 7, lines 49-62: teaches web walker is used for following links and collecting linked documents which these documents can be randomly selected; another process is performed by determining the frequency of linkage for each document within the document collection and on col. 10, lines 15-60: teaches a query to one or more search engines, in other words, each search engine such as a "web crawl" can obtained linked pages and can request other search engine to obtain more linked pages (as a second random walk).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Pitkow to provide a way for randomly selecting linked documents to create a document collection incorporated into a search engine to contain the document collection for other search engines to query to obtain more linked pages which will aid the user in finding the desired information.

Regarding dependent claims 22 and 50, Pitkow discloses:

wherein b) comprises determining a quality metric responsive to the number of documents that link to the target document, and responsive to the quality metric of the linking

documents (Pitkow on col. 10, lines 32-60 and col. 11, lines 10-15: determining "in-link" frequency of a plurality of web pages).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Pitkow to provide a way for randomly selecting linked documents to create a document collection incorporated into a search engine to contain the document collection for other search engines to query to obtain more linked pages which will aid the user in finding the desired information.

21. **Claims 23 and 51 are rejected under 35 U.S.C. 103(a) as being unpatentable over Pitkow, as applied to claims 13-14, 20-22, 26, 41-42, 48-50, 54, and 58-59 above, in view of Page (USPN 6,285,999 B1 – filed on 1/1998).**

Regarding dependent claims 23 and 51, Pitkow discloses the invention substantially as claimed as described *supra*. However, Pitkow does not explicitly disclose "wherein b) comprises determining a value for: $R(p)=d/T+(1-d)\Sigma\ R(pi)/C(pi)$ where: $R(p)$ is the PageRank of target document p; $R(pi)$ is the PageRank of document pi; T is the total number of documents in the document set; d is a damping factor such that $0 < d < 1$; documents p1, , pk each contain at least one link to target document p; and $C(p)$ is the number of links out of document p".

Page on col. 6, lines 12-60 and col. 7, lines 56-64: teaches ranking of a each document in a database containing plurality of linked documents; determining the damping factor; and determining documents containing forward links to determine a value.

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Page into Pitkow to provide a way to rank each document

in a database of a plurality of documents; determine damping factor; and determine documents containing forward links to determine a value, as taught by Page, incorporated into the ranking process of Pitkow in order for the ranking method to provide superior results.

Response to Arguments

22.   Regarding Applicant's arguments filed on 3/17/03 are moot in view of the new ground(s) of rejection as necessitated by amendment.

Regarding Applicant's remarks on pages 52-53, 55-61, and 65-76:

Pitkow does disclose "random walking" on col. 7, lines 15-31 and on col. 10, lines 15-60: teaches a "web walker" and randomly selecting and collecting linked documents; the collection may be gathered based on sampling of web pages, through a "web crawl" where linked pages are obtained, or through a query to one or more search engines; a list of web sites is obtained and further processing of ranking in-links of the list of web pages or web sites is performed.

Pitkow does disclose "a second two-level random walk" on col. 7, lines 49-62: teaches web walker is used for following links and collecting linked documents which these documents can be randomly selected; another process is performed by determining the frequency of linkage for each document within the document collection and on col. 10, lines 15-60: teaches a query to one or more search engines, in other words, each search engine such as a "web crawl" can obtained linked pages and can request other search engine to obtain more linked pages (as a second random walk).
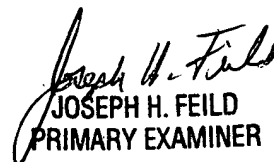
Conclusion

23.   Any inquiry concerning this communication or earlier communications from the examiner should be directed to Almari Romero whose telephone number is (703) 305-5945. The examiner can normally be reached on Mondays - Fridays (7:30am - 4:00pm).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather Herndon can be reached on (703) 308-5186. The fax phone numbers for the organization where this application or proceeding is assigned are (703) 746-7239 for regular communications and (703) 746-7238 for After Final communications.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is (703) 305-4700.

AR
June 12, 2003

JOSEPH H. FEILD
PRIMARY EXAMINER